Integrals of motion for one-dimensional Anderson localized systems

Ranjan Modak[1] , Subroto Mukerjee[1,2], Emil A. Yuzbashyan[3] and B. Sriram Shastry[4],

[1] Department of Physics, Indian Institute of Science, Bangalore 560 012, India
[2] Centre for Quantum Information and Quantum Computing, Indian Institute of Science, Bangalore 560 012, India
[3]Center for Materials Theory, Rutgers University, Piscataway, NJ 08854, USA
[4]Physics Department, University of California, Santa Cruz, CA 95064, USA

Abstract. Anderson localization is known to be inevitable in one dimension for generic disordered models. Since localization leads to Poissonian energy level statistics, we ask if localized systems possess "additional" integrals of motion as well, so as to enhance the analogy with quantum integrable systems. We answer this in the affirmative in the present work. We construct a set of nontrivial integrals of motion for Anderson localized models, in terms of the original creation and annihilation operators. These are found as a power series in the hopping parameter. The recently found Type-1 Hamiltonians, which are known to be quantum integrable in a precise sense, motivate our construction. We note that these models can be viewed as disordered electron models with *infinite-range* hopping, where a similar series truncates at the linear order. We show that despite the infinite range hopping, all states but one are localized. We also study the conservation laws for the disorder free Aubry-Andre model, where the states are either localized or extended, depending on the strength of a coupling constant. We formulate a specific procedure for averaging over disorder, in order to examine the convergence of the power series. Using this procedure in the Aubry-Andre model, we show that integrals of motion given by our construction are well-defined in localized phase, but not so in the extended phase. Finally, we also obtain the integrals of motion for a model with interactions to lowest order in the interaction.

PACS numbers: 02.30.Ik, 05.30.-d, 05.45.Mt, 72.15.Rn

1. Introduction

The simplest theoretical model to study localization for non-interacting particles in the presence of disorder was proposed by Anderson [1]. A single particle localized state has a wavefunction that decays exponentially about some point in space over a characteristic localization length. In three dimensions, localized states exist below a certain energy (the mobility edge) for a given strength of disorder. A disordered electronic system is thus localized if its Fermi energy lies below the mobility edge. In one and two dimensions, an infinitesimal amount of disorder is sufficient to localize all single particle states and thus a disordered non-interacting electronic system is always localized [2, 3].

Recent developments in the area of eigenstate thermalization [4, 5, 6] relate closely to the above well established notions of Anderson localization. In this context, it is believed that an isolated localized eigenstate does not thermalize, in the sense that no subsystem of it can be brought into thermal equilibrium by exchanging heat with the rest of the system. An analogous statement can be made about information, as defined through an appropriate partial trace of the density matrix. A related feature of such a system is the lack of level repulsion in its energy level spectrum. This can be thought of as arising from the presence of almost degenerate states localized so far apart that they are unable to hybridize to lift the degeneracy.

The effect of interactions on such systems is very interesting. Interactions among the elementary degrees of freedom generically tend to drive the system towards thermalization and delocalization [7, 8]. This tendency competes with the the one that causes localization in the presence of disorder. Understanding the resultant phenomenon of many body localization, that is observed for sufficiently strong disorder, is currently a very active area of research [9, 10, 11, 12, 13].

Another class of systems that fail to thermalize are integrable ones. These often contain a variable parameter (such as interaction or external field strength, which we denote here as y) and possess a set of similarly dynamical (i.e. depending on the parameter) integrals of motion. Standard examples of such systems are the one-dimensional Hubbard and XXZ models. In these examples, the integrals of motion I_k are polynomial in y with the order of the polynomial [14, 15, 16, 17, 18, 19, 20, 21] increasing with k. An arbitrary linear superposition of all integrals $Q = \sum_k a_k I_k$ – also an integral in its own right – is an infinite power series in y. Gaudin magnets[22, 23] on the other hand provide examples of integrable models where all conserved charges‡ are linear in y. It should be emphasized that there is no generally accepted precise notion of integrability in quantum mechanics [24, 25] in contrast to classical mechanics where it is unambigous. However, we do not dwell on this issue in present work§. The only aspect important for us here is the existence of parameter-dependent conservation laws.

‡ We use the term conserved charges interchangeably with conservation laws or integrals of motion.
§ By integrable we will generally mean quantum many-body models colloquially recognized as such, see examples in this paragraph. The only exception are type-1 Hamiltonians that stem from a recently proposed well-defined notion of quantum integrability [26, 27, 28, 29, 25, 30].

Conserved charges greatly constrain the dynamics of integrable systems. As a result, when started off from an arbitrary initial state in isolation, these systems do not evolve in a way that causes thermalization in the sense of the above paragraph [6, 31]. Additionally the usual space time symmetries result in degeneracies in the energy level spectrum, and hence a lack of level repulsion [32]. The addition of perturbations destroys such conservation laws and restores level repulsion, although the strength of the perturbations has a non-trivial finite-size dependence [33, 34, 35].

In this context, it is natural to ask in what ways are localized systems similar to integrable ones. In particular we may ask if (parameter dependent) conservation laws, similar to those in integrable systems exist for localized systems. It has been argued in the context of many-body localization that they do, and results related to the growth of entanglement in these systems are predicated on their existence [36, 37, 38]. However, obtaining the structure of the conserved charges directly in terms of microscopic parameters remains a challenge and effective renormalization procedures need to be employed instead [39, 10]. The situation is less complicated in the absence of interactions since the Hamiltonian is that of a single particle. Nevertheless, obtaining the charges systematically and analytically in terms of the microscopic parameters of the Hamiltonian is non-trivial. In this paper we outline the procedure to do so. We also elucidate the connection between localization and conserved charges.

In this work we study a general one dimensional model with on-site disorder that can interpolate between models with long-range hopping and the more standard Anderson-type one. The starting point is a Type-1 Hamiltonian reviewed in Refs. [26, 27, 28, 29, 25, 30]. This was introduced as the most simple model of quantum integrability in finite dimensional spaces. This model has infinite ranged hopping, and as such has no inbuilt metric or length scale. We first show by calculating its Participation ratio (PR) [40, 41] the perhaps surprising result that all states except one are localized. This is done as follows: an eigenstate $|\psi\rangle$ of the Hamiltonian is expanded in a basis of position eigenstates on the lattice as $|\psi\rangle = \sum_k c_k|k\rangle$, where k labels the position eigenstates and c_k are the coefficients in the expansion. The PR for this state is then defined as

$$\mathrm{PR}_\psi = \frac{\left(\sum_k |c_k|^2\right)^2}{\sum_k |c_k|^4}. \tag{1}$$

It is usually understood that $\mathrm{PR}_\psi \sim O(1)$ indicates localization while $\mathrm{PR}_\psi \sim O(N)$ – delocalization, where N is the number of sites. While this definition is valid for a fixed wave function, we may also define the PR at a given energy, as later in the paper, where an averaging over disorder realizations is carried out, at a fixed energy.

The Type-1 model has a known set of conservation laws, which inspire the construction of a generic Anderson-type model having only nearest-neighbor hopping. In 1d it is well known that for this model, all single particle eigenstates are localized for any strength of the disorder. The conserved charges of this model are then constructed by analogy with the Type-1 Hamiltonian. These charges are expressed as a power series in the hopping, whose coefficients we determine by means of an algorithm. We also

show that the series, upon disorder averaging over a "non-resonant" ensemble- defined below, is convergent. This provides numerical evidence that the ensemble chosen and the procedure of averaging the coefficients in the conserved charges over the ensemble is meaningful.

We then turn our focus to a model which contains both localized and delocalized phases (i.e. phases in which *all* single particles states are either localized or delocalized). This is the Aubry-Andre model [42], in which the random potential is replaced by a quasi-periodic one. This allows us to test our criterion for the convergence of the power series and clearly elucidate the connection of the conserved charges to localization. Thus, the convergence (divergence) of the power series representation of conserved charges can indeed be identified with the presence (or absence) of localization and the localization-delocalization transition can be located using the charges. Finally, we investigate the effect of interactions and argue that a power series in the interaction becomes intractable and thus obtain the the conserved charges only to first order in it.

We emphasize that the main feature of our construction is that the conservation laws do not depend explicitly on the wavefunctions of the single particle energy eigenstates. In fact, the recursion relations we obtain for the coefficients of the expansions of the conserved charges are the same for *all* generic one dimensional models. Our approach is thus completely model independent requiring no knowledge of exact solutions or properties of energy eigenfunctions.

Another important aspect of the construction of conservation laws we emphasize here, which has not been discussed before is 'gauge freedom' of a certain kind, defined more precisely later. We show that a judicious choice of gauge can bring out important features of the conserved charges, such as the truncation of their series representation at finite order. These features can be obscured in gauges that arise in constructions of these charges from direct applications of standard methods such as the Rayleigh-Schrödinger series or the locator expansion.

2. Lattice Models

We consider a general Hamiltonian of non-interacting particles hopping on a one dimensional lattice with an on-site potential

$$H = H(y) = \sum_i \epsilon_i n_i - y \sum_{ij} t_{ij} c_i^\dagger c_j, \tag{2}$$

where $c_i^\dagger$ and c_i are fermionic creation and annihilation operators, $n_i = c_i^\dagger c_i$ is the number operator, ϵ_i is the on-site disordered potential, and t_{ij} is the hopping between sites i and j. The parameter y is a real number introduced for convenience, which; it allows us to perform an expansion of the conserved charges in its powers.

Our general strategy to construct construct conserved charges for this models will be to first consider the 'unperturbed' Hamiltonian which only has the on-site potential. The conserved charges for this Hamiltonian are simply the operators n_i, which are

independent and commute with each other and the Hamiltonian. It can also be readily seen that the eigenstates of this Hamiltonian are completely localized on the individual sites. Thus the zeroth order Hamiltonian trivially describes a localized system with conserved charges. We now show that upon introducing the hopping, new conserved charges Q_i appear, which can still be labeled by the site indices i while the system remains localized. To do this, we consider different types of hopping parameters t_{ij}.

3. Type-1 Hamiltonians

We now summarize a known set of conserved charges Q_j. We rework the construction in Refs. [27, 28, 25], in a fashion that suggests a natural generalization for short ranged models. These charges are linear in the hopping (or the parameter y), and commute exactly with the Hamiltonian of the Type-1 family. The Type-1 Hamiltonian is obtained from Eq. (2) by specializing to infinite ranged hopping $t_{ij} = \gamma_i\gamma_j$, with arbitrary parameters γ_j. Specializing to $j = 0$ we write down the charge Q_0

$$Q_0 = n_0 - y\sum_{k\neq 0}\frac{1}{\epsilon_0-\epsilon_k}[t_{0k}(c_0^\dagger c_k + c_k^\dagger c_0)\underbrace{-\alpha_k^0 n_0 - \beta_k^0 n_k}], \tag{3}$$

where α_k^0 and β_k^0 are yet to be determined.The commutator of Q_0 and H vanishes to linear order in y by construction.The surviving term is of $O(y^2)$ and is given by

$$[Q_0, H] = y^2\sum_{jk}\frac{1}{\epsilon_0-\epsilon_k}\quad [A(0,j,k)(c_0^\dagger c_j - c_j^\dagger c_0) + B(0,j,k)(c_k^\dagger c_j - c_j^\dagger c_k)] = 0, \tag{4}$$

where

$$\begin{aligned} A(0,j,k) &= t_{0k}t_{kj} - \alpha_k^0 t_{0j}\\ B(0,j,k) &= t_{0j}t_{0k} - \beta_k^0 t_{kj}. \end{aligned} \tag{5}$$

A few words on the form of Eq. (3) are appropriate here. The last two terms $-\alpha_k^0 n_k - \beta_k^0 n_0$ commute with $H(y=0)$ trivially, since they are expressed in terms of the number operators. These actually represent a particularly convenient "gauge choice", their presence enables the second order term $O(y^2)$ to vanish, and thus the commutator series to truncate exactly for the Type-1 matrices. The requirement that $[Q_0, H] = 0$ is satisfied by the following form of t_{ij}.

$$\begin{aligned} t_{ij} &= \gamma_i\gamma_j\\ \alpha_j^0 &= \gamma_j^2\\ \beta_j^0 &= \gamma_0^2, \end{aligned} \tag{6}$$

this gives $A(0,j,k) = B(0,j,k) = 0$. It is straightforward to extend this definition to arbitrary Q_j, and further to show that $[Q_i, Q_j] = 0\ \forall i,j$, so the operators Q_i are indeed

the conserved charges of the Hamiltonian H[26, 27]. The Hamiltonians described by t_{ij} of the form given in Eq. (6) are called Type 1 [27, 25], and can also be interpreted geometrically as representing a 'd-simplex' [43].

3.1. PR for Type-1 Hamiltonians

All single particle states of Type-1 Hamiltonians (6), except possibly the ground state for $y > 0$ or the highest energy state for $y < 0$ are localized, see e.g. Fig. 1.

This can be understood in more detail from the exact solution for the spectrum of these models[27]. Exact un-normalized single particle eigenstates of the Hamiltonian (6) are

$$|E\rangle = \sum_{i=0}^{N-1} \frac{\gamma_i c_i^\dagger}{E - \epsilon_i}|0\rangle, \tag{7}$$

and the corresponding eigenvalues E (energies) are solutions of the equation

$$\sum_{i=0}^{N-1} \frac{\gamma_i^2}{E - \epsilon_i} = -\frac{1}{y}. \tag{8}$$

Suppose ϵ_i are ordered in the ascending order. By plotting the left hand side of Eq. (8) as a function of E, one can verify that it has $N-1$ real roots $E_1, E_2, \ldots E_{N-1}$ located between consecutive ϵ_i, i.e. $\epsilon_{i-1} < E_i < \epsilon_i$. The remaining root E_0 is also real and is below ϵ_0 (ground state) for $y > 0$ and above ϵ_{N-1} for $y < 0$ (highest excited state).

Eqs. (7) and (8) also provide an exact solution for one fermion (Cooper) pair and one spin flip sectors of the BCS and Gaudin models, respectively,

$$\begin{aligned} H_{\rm BCS} &= \sum_{i,\sigma=\uparrow,\downarrow} \epsilon_i c_{i\sigma}^\dagger c_{i\sigma} - y \sum_{ij} c_{i\downarrow}^\dagger c_{i\uparrow}^\dagger c_{j\uparrow} c_{j\downarrow}, \\ H_i(x) &= s_i^z - y \sum_{j\neq i} \frac{\vec{s}_i \cdot \vec{s}_j}{\epsilon_i - \epsilon_j}, \end{aligned} \tag{9}$$

where $c_{i\sigma}$ are spin-full fermions and $\vec{s}_i$ are quantum spins of arbitrary magnitudes s_i, see Ref. [27] for details. For the BCS [Gaudin] model one needs to replace $c_i^\dagger \to c_{i\downarrow}^\dagger c_{i\uparrow}^\dagger$ [s_i^+] in Eq. (7), set $\gamma_i = 1$ [$\sqrt{2s_i}$] and the corresponding eigenvalue is equal to $2E$ [$2s_i(E-\epsilon_i)^{-1}$] rather than E. Our results for the PR of Type-1 Hamiltonians therefore also apply to these sectors of these models.

The PR defined through Eq. (14) reads

$$\mathrm{PR}_E = \frac{\left[\sum_i \frac{\gamma_i^2}{(E-\epsilon_i)^2}\right]^2}{\sum_i \frac{\gamma_i^4}{(E-\epsilon_i)^4}}. \tag{10}$$

For concreteness we take $y > 0$. Then, the ground state is $E_0 < \epsilon_0$. We assume that most γ_i are of the same order of magnitude and consequently the vector with components γ_i

is delocalized. Further, we take ϵ_i to lie in a fixed interval that does not scale with N, e.g. from $-w$ to w.

For excited states E_k is between ϵ_{k-1} and ϵ_k. The summations in the numerator and denominator of Eq. (10) both come from ϵ_i in a small vicinity of ϵ_k for large N and converge as $\sum_n n^{-2}$ and $\sum_n n^{-4}$, respectively, where $n = |i-k|$. The numerator and the denominator scale as $[\gamma_k^2/\delta^2]^2$ and γ_i^4/δ^4, where $\delta \propto 1/N$ is the mean level spacing between ϵ_i in the vicinity of ϵ_k. Therefore, PR_{E_k} is of order 1 (much smaller than N) meaning excited states are always localized. Fig. 1 shows PR for $N = 10^3$ uncorrelated random ϵ_i uniformly drawn from an interval $(-1, 1)$ and the same distribution of γ_i.

Consistent with our numerical results, we estimate the largest PR for excited states to scale as $\ln N$, i.e.

$$\mathrm{PR}_{E_k}^{\max} \approx \alpha \ln N, \tag{11}$$

for large N, where α depends on N much weaker than $\ln N$. Such values of PR come from clustering in ϵ_i. Indeed, suppose spacings $\delta_i = \epsilon_{i+1} - \epsilon_i$ between m of ϵ_i for i from k to $k+m$ are all much smaller than δ_{k-1} and, moreover, $\epsilon_{k+m} - \epsilon_k \leq \delta_{k-1}$. It follows from Eq. (10) that $\mathrm{PR}_{E_k} \approx \mathrm{PR}_{E_{k+m+1}} \approx m$ because the above ϵ_i contribute most to these PRs. Normalized spacings $s_i = \delta_i/\delta$ are distributed according to the Poisson distribution $P(s)ds = e^{-s}ds$. The probability of having m spacings between 0 and $s_0 \ll 1$ is then roughly s_0^m. We need $ms_0 \leq 1$ and also $Ns_0^m = 1$ so that at least one such clustering occurs‖ . This implies $m \approx \ln N / \ln(\ln N)$ and Eq. (11) follows. Numerically we find that typical values of $\alpha \approx 1-3$ and averaged over disorder $\bar{\alpha} \approx 1.7$, at least for $N = 2^4 - 2^{12}$. Note that according to this argument such large values of PR typically come in pairs spaced by $m+1$, roughly equal to the value of the PR itself. We also stress that, in contrast to the largest PR, a typical (and average) PR is something between one and three for any N (does not scale) as can be seen from Fig. 1.

It is interesting to compare this $\ln N$ behavior to the flat band localization studied earlier [44, 45]. The latter leads to a (weakly) divergent PR in the localized regime, a phenomenon that is viewed as corresponding to critical (power law type) localization. The Type-1 Hamiltonian kinetic energy may also be viewed as a "flat band" model, with a flat dispersion for all except one state. Indeed, for $t_{ij} = \gamma_i\gamma_j$ all but one eigenvalues of the second term in Eq. (2) are zero. The non-zero eigenvalue (ground state for $y > 0$) corresponds to the eigenstate $\gamma_i c_i^\dagger|0\rangle$.

Let us consider limits $y \to 0$ and $y \to \infty$ separately. When $y \to 0$ all states are localized as expected. Indeed, Eq. (8) implies $E_k \to \epsilon_k$, summations in Eq. (10) are dominated by the $i = k$ term and we obtain $\mathrm{PR}_{E_k} = 1$ for all k. When $y \to \infty$ excited states are localized as before because E_k for $k \geq 1$ remains trapped in the interval $(\epsilon_{k-1}, \epsilon_k)$. The ground state energy on the other hand diverges – Eq. (8) implies

‖ More precisely, the probability that m of ϵ_i occur in an interval of length δ for Poisson distribution is $e^{-1}/m!$, which however still leads to the same estimate (11)

$E_0 \to -y\sum_i \gamma_i^2$. Then, ϵ_i are negligible as compared to E_0 in Eq. (10) and

$$\mathrm{PR}_{E_0} = \frac{\left[\sum_i \gamma_i^2\right]^2}{\sum_i \gamma_i^4}, \tag{12}$$

which is of order N according to our choice of γ_i. The ground state is therefore delocalized for $y \to \infty$. It undergoes a localization-delocalization crossover at a certain y_c, which we estimate below in this section.

It is possible to evaluate the PR analytically to leading order in $1/N$ for distributions of ϵ_i and γ_i with negligible short range fluctuations (such that the spacing $\delta_i = \epsilon_{i+1} - \epsilon_i$ changes slowly with i – $|\delta_{i+1} - \delta_i|/\delta_i$ is of order $1/N$ for all i – and similarly for γ_i). For simplicity, let us take constant γ_i, which we can set to one with no loss of generality, and equally spaced ϵ_i, i.e. $\delta_i = \delta = 2w/N$.

For excited states, we write $E_k = \epsilon_k - \alpha_k\delta$, where $0 < \alpha_k < 1$, and solve Eq. (8) for α_k to the leading order in $1/N$ as described in Appendix B of[46]. This yields

$$\cot \pi\alpha_k = \frac{\delta}{\pi y} + \frac{1}{\pi}\ln\frac{\epsilon_k + w}{w - \epsilon_k} \equiv f(\epsilon_k). \tag{13}$$

We note that $\lambda = y/\delta$ is the proper dimensionless coupling constant in the sense that it must stay finite in the $N \to \infty$ limit. This is because the second summation in Eq. (2) scales as N^2 for $t_{ij} = \gamma_i\gamma_j$ and our choice of γ_i. Therefore, we need $y \propto \delta \propto 1/N$ so that both terms in Eq. (2) are extensive in the thermodynamic limit. For the BCS Hamiltonian in Eq. (9), so defined λ is the dimensionless superconducting coupling[47].

Eq. (10) becomes to leading order in $1/N$

$$\mathrm{PR}_{E_k} = \frac{\left[\sum_{n=0}^{\infty}\left(\frac{1}{(n+\alpha_k)^2} + \frac{1}{(n+1-\alpha_k)^2}\right)\right]^2}{\sum_{n=0}^{\infty}\left(\frac{1}{(n+\alpha_k)^4} + \frac{1}{(n+1-\alpha_k)^4}\right)}, \tag{14}$$

which evaluates to

$$\mathrm{PR}_{E_k} = \frac{3}{1 + 2\cos^2\pi\alpha_k} = \frac{3 + 3f^2(\epsilon_k)}{1 + 3f^2(\epsilon_k)}. \tag{15}$$

This answer is in good agreement with numerics already for $N = 20$, see also Fig. 1. Note that $1 \le \mathrm{PR}_{E_k} \le 3$.

We saw above that the ground state energy $E_0 \to -\infty$ as $y \to \infty$, while $E_0 \to \epsilon_0$ for $y \to 0$. Let y be large enough that E_0 is well separated from ϵ_0. Then, we can replace summation in Eq. (8) with integration and obtain

$$\ln\frac{E_0 - w}{E_0 + w} = \frac{\delta}{y} = \frac{2w}{Ny}. \tag{16}$$

Performing the same replacement in Eq. (10) and using Eq. (16), we derive

$$\mathrm{PR}_{E_0} = \frac{3N}{1 + 2\cosh(\delta/y)}. \tag{17}$$

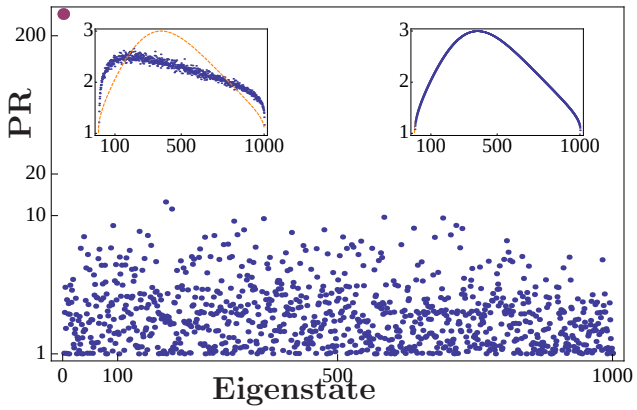

Figure 1. PR of eigenstates of a Type-1 Hamiltonian for $y = .004$, $N = 10^3$ in ascending order according to the energy. Each ϵ_i and γ_i is an independent random number uniformly distributed in an interval $(-1, 1)$. Larger circle near the left top corner indicates the ground state, which is extended. Left inset is the same as above, but averaged over 10^3 realizations of disorder and compared to Eq. (15) for the same y, N, w. The right inset shows the PR (except the ground state) for $N = 10^3$ equally spaced ϵ_i, $\gamma_i = 1$ and $y = .004$ similarly compared to Eq. (15) (the two curves are indistinguishable).

Note that in the limit $y \to \infty$, $\mathrm{PR}_{E_0} = N$ in agreement with Eq. (12). This expression also allows us to estimate the value y_c beyond which the ground state becomes extended. We obtain $\lambda_c = y_c/\delta \approx 1/\ln N$. This also corresponds to the coupling for which the gap in the spectrum $\Delta = E_1 - E_0 \approx -w - E_0$ becomes comparable to the spacing δ. For a superconductor described by the BCS model (9) this localized-extended crossover translates into a normal-superconducting one[48, 49]. As $N \to \infty$ this crossover becomes a quantum phase transition at $\lambda = 0$, i.e. any infinitesimal coupling is sufficient to make the ground state extended (superconducting). The localized character of the excited states for the specific case of $\gamma_i = 1$ has been demonstrated in a previous work as well [43].

4. A model with finite-ranged hopping

We now consider the following Anderson-type model in one dimension with nearest neighbor hopping.

$$\begin{aligned} H &= \sum_i \epsilon_i n_i - yt \sum_i (c_i^\dagger c_{i+1} + h.c.) \\ &= H_0 + yH_1. \end{aligned} \tag{18}$$

This corresponds to the case with $t_{ij} = t$ for $|i - j| = 1$ and 0 otherwise for the general Hamiltonian in Eq. (2). H_0 is the zeroth order Hamiltonian with only the on-site potential and H_1 contains the hopping. It is known that all single particle eigenstates of this Hamiltonian are localized [1, 3].

4.1. Construction of the conserved charges

Proceeding as for the case of Type-1 Hamiltonians, we focus on the conserved charge Q_0, corresponding to the site $i = 0$, which to lowest order is equal to n_0. However, in this case Q_0 is not simply linear in y. In fact, it can be argued that the an expansion of Q_0 in the hopping does not truncate at any finite order in the thermodynamic limit. Indeed, as explained in the Introduction, conserved charges are generally infinite power series in y. We thus assume Q_i of the form

$$Q_i = P_{i0} + yP_{i1} + y^2 P_{i2} + \dots, \tag{19}$$

where $P_{i0} = n_0$ and $P_{i1}, P_{i2} \dots$ are operators to be determined in terms of the microscopic parameters subject to the condition $[Q_i, H] = 0$. For concreteness, we first take our one dimensional system to be a finite-sized ring of $N + 1$ sites going from 0 to N.

Since the Hamiltonian H and and all the zero order charges n_i are quadratic in the creation and annihilation operators, we take all the operators $P_{i1}, P_{i2} \dots$ to be similarly quadratic, i.e.

$$P_{im} = \sum_{jk} \eta^{(m)}_{jk}(i) c^\dagger_k c_j, \tag{20}$$

where the symmetric coefficients $\eta^{(m)}_{jk}(i) = \eta^{(m)}_{kj}(i)$ are to be determined. We have

$$[Q_i, H] = [P_{i0}, H_0] + \sum_m y^{m+1} \left([P_{im}, H_1] + [P_{im+1}, H_0] \right).$$

The requirement that the commutator vanishes to all orders in y requires

$$[P_{im}, H_1] + [P_{im+1}, H_0] = 0, \tag{21}$$

and yields a recursion relation among $\eta's$.

$$\eta^{(m+1)}_{ab}(i) = \delta_{ab} R^{(m+1)}_a(i) + \frac{1 - \delta_{ab}}{\epsilon_a - \epsilon_b} \sum_j [(t_{aj} \eta^{(m)}_{jb}(i) - \eta^{(m)}_{aj}(i) t_{jb}], \tag{22}$$

with initial conditions $\eta^{(0)}_{ab}(i) = \delta_{ia}\delta_{ib}$. The diagonal term $R^{(m+1)}_a$ represents a "gauge" freedom, since the corresponding term in P_{im} commutes trivially with H_0. We further discuss this freedom below. Specializing to the case of nearest neighbor hopping Eq. (18) and with $i = 0$, it can be verified that terms present in P_{0m} are of the form

- $\eta^{(m)}_{0m}(c^\dagger_0 c_m + c^\dagger_m c_0)$
- $\eta^{(m)}_{0,N-(m-1)}[c^\dagger_0 c_{N-(m-1)} + c^\dagger_{N-(m-1)} c_0]$
- $\sum_{ij \neq m, |i-j| = \text{even} \leq m} \eta^{(m)}_{ij}(c^\dagger_i c_j + c^\dagger_j c_i)$ (if m is even)
- $\sum_{ij \neq m, |i-j| = \text{odd} \leq m} \eta^{(m)}_{ij}(c^\dagger_i c_j + c^\dagger_j c_i)$ (if m is odd)

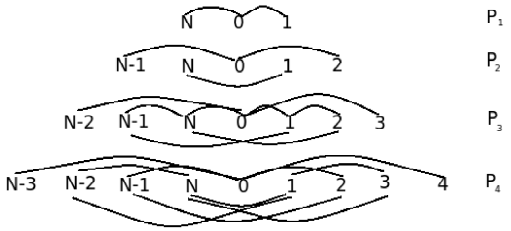

Figure 2. Schematic diagram showing hopping terms present in the operators $P_1 - P_4$. The base site 0 is in the middle and its neighbors are sites 1 and N, since we imposed periodic boundary conditions. Lines connecting pairs of sites indicate the presence of the corresponding hopping term in the operator P_m. Note that the range of the hopping in P_m increases with m.

This is shown schematically in Fig. 2 for the first few P_m.

The $Q_i's$ are related to each other by translating all site indices in the above relations by an appropriate number. By construction, they all commute with H. Since H is generally non-degenerate, this implies Q_i also commute among themselves, $[Q_i, Q_j] = 0$ $\forall\ i, j$. To see this, first recall that for Hermitian matrices $[A, B] = [A, C] = 0$ implies $[B, C] = 0$ as long as eigenvalues of A are non-degenerate. All operators involved in the above construction of Q_i are of the form $\hat{A} = \sum_{ij} A_{ij} c_i^\dagger c_j$, where A_{ij} is a Hermitian $N \times N$ matrix, which represents operator $\hat{A}$ in the sector with total particle number $n = 1$. Moreover, the commutativity of any two such operators is equivalent to that of the underlying matrices. Eigenvalues of the Hamiltonian in the $n = 1$ sector at $y = 0$ are ϵ_i, which are assumed to be distinct, i.e. the corresponding matrix is non-degenerate at $y = 0$. By continuity of the eigenvalues in y, it remains non-degenerate in some finite interval (until the first level crossing) of the real axis containing $y = 0$. Thus, $[Q_i, Q_j] = 0\ \forall\ i, j$ in this interval of y. But, as can be seen e.g. from the above construction of Q_0, commutativity of Q_i on any finite interval of values of y implies that they commute for all y.

We noted above that commutation relations (21) and consequently recursion relations (22) do not constrain the diagonal part of the coefficients $\eta^{(m)}$, i.e. $R_r^{(m)}$, for $m \geq 1$. The choice of $R_a^{(m)}(i)$ however *does* affect the off-diagonal part of $\eta^{(k)}$ for $k > m$. In our construction of Q_i we set $R_a^{(m)}(i) = 0$ for all $m \geq 1$, since this leads to the most compact description of these objects. we will refer to this as the *standard gauge*. Conserved charges $\widetilde{Q}_i$ resulting from *any other choice* $R_a^{(m)}(i)$ uniquely relate to our standard gauge Q_i's, a brief calculation shows their relationship is

$$\widetilde{Q}_i = Q_i + \sum_m y^m \sum_r R_r^{(m)}(i) Q_r. \tag{23}$$

Another advantage of our choice of a gauge is in a simple relationship between the

Hamiltonian (18) and the conserved charges, namely,

$$H = \sum_i \epsilon_i Q_i. \tag{24}$$

To see this, consider the difference

$$H - \sum_i \epsilon_i Q_i = yW_1 + y^2 W_2 + y^3 W_3 + \ldots, \tag{25}$$

where W_i are y-independent operators. Note that the zeroth order term cancels in the difference. Since H commutes with all Q_i, the right hand side (RHS) of Eq. (25) must also commute. This implies in particular $[W_1, n_i] = 0$ for all i (from the coefficient at the lowest power of y in the commutator of the RHS with Q_i), which in turn means that $W_1 = \sum_i r_i^1 n_i$. Now note that the left hand side (LHS) has zero diagonal matrix elements, i.e. no terms of the form $c_r^\dagger c_r$. This is because the zeroth order term is absent, while higher order terms have no diagonal matrix elements since $\eta_{rr}^{(m)} = 0$ for all $m \geq 1$ in our gauge (and similarly the diagonal is absent in other Q_i). Then, the diagonal matrix elements must vanish on the RHS as well, to all orders in y. In particular, $r_i^1 = 0$, i.e. $W_1 = 0$ and

$$H - \sum_i \epsilon_i Q_i = y^2 W_2 + y^3 W_3 + \ldots \tag{26}$$

Applying the same argument to the RHS of this equation we similarly obtain $W_2 = 0$ etc., until we finally arrive at Eq. (24).

4.2. Type-1 Hamiltonians redux

We have seen above that the conserved charges are power series in the hopping. This is unlike the case of Type-1 Hamiltonians, where the power series *truncates* after the first term. The gauge where the series truncates corresponds to having distinct terms for $m = 1$, one can see in Eq. (3) (the gauge terms are indicated in the lower braces).

It is an amusing exercise to determine the correct gauge terms that lead to truncation, starting from the recursion relations Eq. (22). To obtain Type-1 Hamiltonians we set $t_{ij} = \gamma_i \gamma_j$, so that the recursions simplify to

$$\begin{aligned} \eta_{ab}^{(m+1)}(i) &= \delta_{ab} R_a^{(m+1)}(i) - \frac{1-\delta_{ab}}{\epsilon_a - \epsilon_b}\left(Y_{ab}^{(m)}(i) - Y_{ba}^{(m)}(i)\right) \\ Y_{ab}^{(m)}(i) &= \sum_j \eta_{aj}^{(m)}(i)\gamma_j\, \gamma_b. \end{aligned} \tag{27}$$

With the initial condition $\eta_{ab}^{(0)}(i) = \delta_{ia}\delta_{ib}$, we obtain at the first level

$$\eta_{ab}^{(1)}(i) = \delta_{ab} R_a^{(1)}(i) + \frac{\gamma_a \gamma_b (1-\delta_{ab})}{\epsilon_a - \epsilon_b}\left(\delta_{ib} - \delta_{ia}\right). \tag{28}$$

At this point we pause and ask if we can choose the gauge term $R_a^{(1)}(i)$ such that $\eta_{ab}^{(2)}(i)$ can be made to vanish identically, so that the iterations stop at the first level. From

Eq. (27) we see that the relevant condition is the vanishing of $\left(Y^{(1)}_{ab}(i) - Y^{(1)}_{ba}(i)\right)$. Using Eq. (27) compute

$$Y^{(1)}_{ab}(i) = \gamma_a\gamma_b\{R^{(1)}_a(i) + \frac{\gamma_i^2}{\epsilon_a - \epsilon_i}(1-\delta_{ia}) - \delta_{ia}\sum_j{}' \frac{\gamma_j^2}{\epsilon_j - \epsilon_i}\}. \tag{29}$$

We may choose $R^{(1)}$ so that the term in braces vanishes, thus leading to the truncation of the iterations. From Eq. (28) we have the complete first order term, and we can proceed to construct the charge (denoting the currents by the symbol $\widetilde{Q}$)

$$\widetilde{Q}_i = n_i + y\sum_{ab}\eta^{(1)}_{ab}(i)c^\dagger_a c_b, \tag{30}$$

which is identical to that in Eq. (3).

The use of the gauge term here is very special, and guided by our understanding of this model. On the other hand, we could by default set all the gauge terms $R^{(m)}$ to zero, giving us the irreducible (i.e. standard gauge) currents. These no longer truncate even for Type-1 Hamiltonians. For completeness we note the second order term for the current in this (standard) gauge

$$\begin{aligned} Q_i = n_i + (y + y^2\sum_j{}' \frac{\gamma_j^2}{\epsilon_j - \epsilon_i}) \times \sum_{i\neq j}\frac{\gamma_i\gamma_j}{\epsilon_j - \epsilon_i}(c^\dagger_i c_j + c^\dagger_j c_i) \\ + y^2\gamma_i^2 \times \sum_{a,b}{}' \frac{\gamma_a\gamma_b}{(\epsilon_a - \epsilon_i)(\epsilon_a - \epsilon_j)}c^\dagger_i c_j + O(y^3). \end{aligned} \tag{31}$$

Thus the Type-1 Hamiltonians allow for variety of expressions of the constants of motion. To establish their equivalence in general is a subtle problem, where some surprising results have been found quite recently in Ref. [30].

This type of gauge choice, made explicit in our construction could be exploited further to test the possibility that the series can take simpler forms, as compared to a brute force expansions to infinite order. We leave this interesting question for future investigation.

4.3. Currents found from the Rayleigh Schrödinger (locator) expansion

A natural question that arises is the relationship between the currents found above and those found from a brute force expansion of the projection operators of the Anderson model in powers of the coupling constant y. The model has a formal single particle eigenfunction expansion in the form

$$|\Psi(y)\rangle = \sum_k u_{0k}(y)c^\dagger_k|0\rangle, \tag{32}$$

with an initial condition localized say at the site 0 as $u_{0k}(0) = \delta_{k0}$. The projector $Q = |\Psi(y)\rangle\langle\Psi(y)|$ can be expanded in a series in y

$$\hat{Q} = \sum_{j,k} u_{0j}u^*_{0k}c^\dagger_j c_k = \hat{P}(0) + y\hat{P}(1) + y^2\hat{P}(2) + \ldots \tag{33}$$

so that the basic expansion of the wave functions in a Rayleigh Schrödinger (RS) series in y generates the conserved currents. We can use the standard result in text books ¶ to write a perturbative expansion for the state at site 0 with standard normalization to $u_{00} = 1$ as

$$|\Psi(y)\rangle = c_0^\dagger|0\rangle + \sum_{k\neq 0} u_{0k} c_k^\dagger|0\rangle, \tag{34}$$

with a power series expansion for u_{0k}

$$u_{0k} = -y\frac{t_{0k}}{\epsilon_0 - \epsilon_k} + y^2 \sum_{l\neq 0} \frac{t_{kl}t_{l0}}{(\epsilon_0 - \epsilon_k)(\epsilon_0 - \epsilon_l)} - y^2 \frac{t_{00}t_{k0}}{(\epsilon_0 - \epsilon_k)^2} + O(y^3). \tag{35}$$

Using this expansion, we may generate the series Eq. (33), the result is explicitly stated below in Eq. (38). From this series we can verify to second order, that this series differs from that in the standard gauge Eq. (19) by specific gauge terms. The advantage of Eq. (19) is that this gauge invariance is manifest in the construction by the nested commutators. On the other hand, Eq. (33-35), corresponds to a particular gauge picked out by the R-S method, and the currents found here are some linear combinations of the ones in Eq. (19) as in Eq. (23).

It seems to us that the series in Eq. (19) possesses an essential simplicity relative to the Rayleigh-Schrödinger series Eqs. (33-35). The R-S perturbation expansion simultaneously determines the energy eigenvalue, and for this purpose very specific gauge terms are needed. On the other hand all terms in Eq. (19) are generated by completely off diagonal terms, those terms that avoid multiple visits to any site. This leads to simpler recursion relations, as in Eq. (22), relative to the RS series. For this reason our numerical work in this paper uses the series in Eq. (19).

4.4. Locator expansion for Type-1 Hamiltonians

The Rayleigh-Schrödinger series can also be constructed for Type-1 Hamiltonians using the exact eigenstates $|E\rangle$ with eigenvalues E as given in Eq. (7) and Eq. (8). The projector $|E\rangle\langle E| = \sum_{ij} \frac{\gamma_i\gamma_j c_i^\dagger c_j}{(E-\epsilon_i)(E-\epsilon_j)}$ can be expanded in y as shown in Eq. (33). In the limit $y \to 0$, the roots of Eq. (8) tend to ϵ_i. We take the root $E \to \epsilon_0$ to obtain $\widetilde{Q}_0$ the conserved charge corresponding to site 0 calculated using the Rayleigh-Schrödinger gauge. Other roots yield other $\widetilde{Q}_i$. Expanding Eq. (8) for E in y near $E = \epsilon_0$, we get

$$E = \epsilon_0 - y\gamma_0^2 + y^2\gamma_0^2 \sum_{i\neq 0} \frac{\gamma_i^2}{\epsilon_0 - \epsilon_i} + O(y^3). \tag{36}$$

¶ For e.g. see Eq. (5.1.44) *Modern Quantum Mechanics*, J J Sakurai (Pearson Education 1994))

Since, the projector diverges in $y \to 0$ limit, we define our conserved charge as $\widetilde{Q}_0 = \frac{(E-\epsilon_0)^2}{\gamma_0^2}|E\rangle\langle E|$ to make it well behaved. $\widetilde{Q}_0$ is given by,

$$\widetilde{Q}_0 = n_0 + \frac{E-\epsilon_0}{\gamma_0}\sum_{j\neq 0}\frac{\gamma_j(c_0^\dagger c_j + c_j^\dagger c_0)}{E-\epsilon_j} + \frac{(E-\epsilon_0)^2}{\gamma_0^2}\sum_{i,j\neq 0}\frac{\gamma_i\gamma_j c_i^\dagger c_j}{(E-\epsilon_i)(E-\epsilon_j)} \tag{37}$$

Then, replacing $\frac{(E-\epsilon_0)^2}{\gamma_0^2} \to y^2\gamma_0^2$ and then $E \to \epsilon_0$, we have obtained $\widetilde{Q}_0$ as a combination of Q_i (see Eq. 3) as follows,

$$\widetilde{Q}_0 = Q_0 - y\sum_{i\neq 0}\frac{\gamma_0^2 Q_i + \gamma_i^2 Q_0}{\epsilon_0 - \epsilon_i} + O(y^3) \tag{38}$$

Other $\widetilde{Q}_k$ can be obtained with the replacement $0 \to k$. Unlike Q_k, there is no indication of the series truncating at any finite order for $\widetilde{Q}_k$.

4.5. Convergence of the power series

The conserved charges constructed above depend on the microscopic parameters of the Hamiltonian, i.e. the hopping and on-site energies. As we shall show later, the same Hamiltonian can have a localized and delocalized phase depending on the values of these parameters. It is thus important to understand if and how the conserved charges themselves differ in the two phases. More precisely, how do the conservation laws "know" whether a particular choice of microscopic parameters produces a localized or delocalized phase?

The answer has to do with their convergence since they are expressed as power series in the microscopic parameters and particle operators. We thus need to state in what sense the power series are convergent. A reasonable condition for convergence is a sufficiently rapid decay of the coefficients η_{ij}^m with increasing m. However, this is complicated by the fact that there are energy difference denominators in the coefficients η_{ij}^m that can cause them to blow up when the on-site energies at two different sites are equal. To avoid this, we restrict ourselves to a particular type of disorder that may be termed "non-resonant". By this we mean any ensemble of ϵ_i, which shows "level repulsion", i.e. the probability of finding ϵ_i very close to each other is very small.

From the random matrix theory, we know that the eigenvalues of a generic matrix display level repulsion in their eigenvalues of various degree, the Gaussian Orthogonal Ensemble (GOE) [50] of real symmetric matrices has the least level repulsion. This condition ensures that perturbative resonances from small denominators, that would otherwise cause individual terms in the expansions of the conserved charges to diverge, are prohibited. This choice is similar to the one involving limited level attraction recently adopted in the context of many-body localization [51].

We have verified that this distribution of onsite energies gives us localization (as indicated from a calculation of the participation ratio) immediately upon switching on

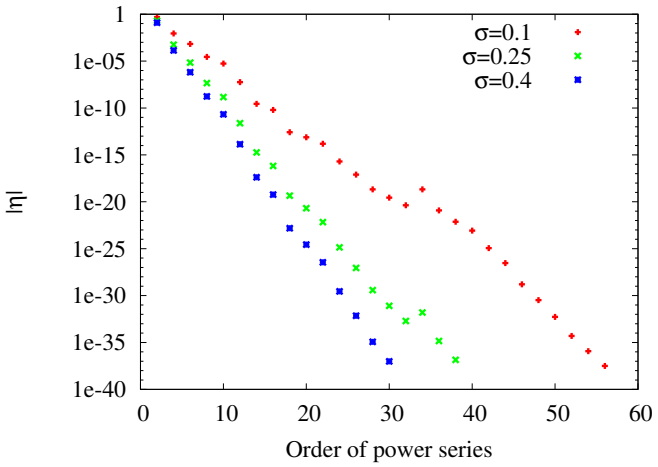

Figure 3. Plot indicating convergence of conserved charges [see Eqs. (19) and (20)] of the Anderson model (18) for $N = 500$. $|\eta|$ represents a typical m-th coefficient η^m_{ij} averaged over a distribution of the on-site disorder ϵ_i, see the end of Sect. 4.5. The plot shows the logarithm of the average as a function of m. ϵ_i are drawn from the eigenvalues of real symmetric matrices whose elements are Gaussian random variables of variance σ =0.1, 0.25 and 0.4, and we set $yt = 1$.

the hopping term. Thus, this particular choice of onsite energies, which is of great convenience from the point of view of calculations, is also not unphysical. The on-site energies ϵ_i are drawn from the eigenvalues of a real symmetric matrices whose elements are taken from a Gaussian random distribution with fixed variance. The eigenvalues of these matrices are assigned randomly to different sites. Different random assignments then constitute different realizations of disorder, which can then be averaged over to check for convergence. The result of this procedure is shown in Fig. 3, where ϵ_i are drawn from the eigenvalues of real symmetric matrices whose elements are taken from a Gaussian distribution of variance σ=0.1, 0.25 and 0.4. It can be seen that the η^m decrease rapidly with increasing order of power series m indicating convergence. We have also checked the convergence of the power series for ϵ_i drawn from the eigenvalues of non-integrable $t - t' - V$ model, which also follow a GOE distribution [33, 52].

Since η^m_{ij} contain more than one term for each m, we checked the convergence of a typical term, which is of the form $\frac{t^m}{(\epsilon_{a_1}-\epsilon_{b_1})(\epsilon_{a_2}-\epsilon_{b_2})....(\epsilon_{a_m}-\epsilon_{b_m})}$. Recall that the m^{th} order term in the calculation of $\tilde{Q}_0$ involves sites with labels between $N - (m - 1)$ and m as can be seen from Fig. 2. Thus, the only values of ϵ_i involved are are those chosen from $[\epsilon_{N-(m-1)}, \epsilon_m]$ (ϵ_0 is at the center) such that $\epsilon_{a_i} \neq \epsilon_{b_i}$, $\forall i$ and $\max |a_i - b_i| = m$

As the aim of this work is to construct conserved charges in localized systems, it is legitimate to ask whether this slightly non-standard choice of disorder distribution produces localization. We have verified this through numerical exact diagonalization by calculating the PR. We find that the PR for different eigenstates is indeed close to zero for systems of size $N = 500$ as shown in Fig. 4, consistent with localization. We thus conclude that our model with on-site energies taken from a GOE distribution does indeed

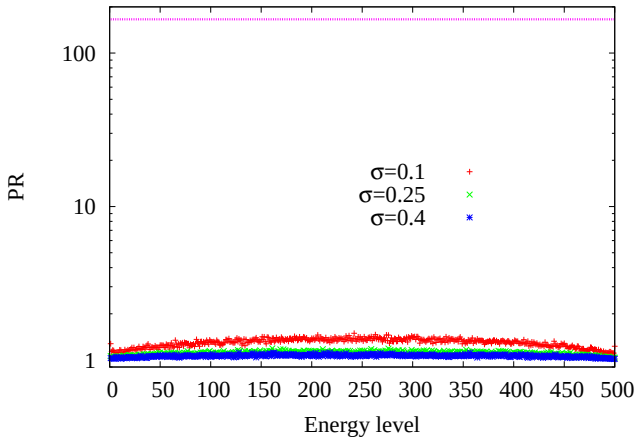

Figure 4. PR of eigenstates of the Anderson model (18) for $N = 500$ numbered in ascending order according to the energy levels. On-site disorder ϵ_i is drawn from the eigenvalues of real symmetric matrices whose elements are Gaussian random variables of variance σ =0.1, 0.25 and 0.4, and we set $yt = 1$. Blue dashed line corresponds to the typical value of PR in delocalized phase.

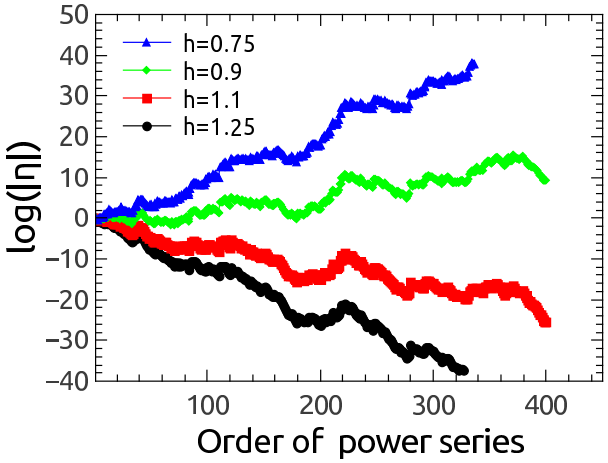

Figure 5. Conserved charges for Aubrey-Andre model converge for $h > 1$ (localized phase) and diverge for $h < 1$ (delocalized phase). Here $|\eta|$ represents a typical m-th coefficient η_{ij}^m in Eq. (20) [see the end of Sect. 4.5], $N = 900$ and $\beta = \frac{\sqrt{5}-1}{2}$. The plot shows $\log|\eta|$ as a function of m.

produce a localized phase. A similar exercise to construct the conservation laws for the above model has been carried out in Ref. [53]. In that work too, the conserved charges have been constructed as infinite operator series but whose coefficients correspond to the amplitudes of a particle to be on the sites of a square lattice whose sides are the physical one dimensional lattice. The recursion relation obtained is between conserved charges on different sites and the convergence of the series is assumed to follow from the exponential decay of the eigenfunctions of the Hamiltonian. In our calculations, we construct the conserved charges directly in terms of the microscopic parameters of the

Hamiltonian and our convergence criterion is not based on any assumption about the nature of the eigenstates of the Hamiltonian. In fact, as we show in the next section, the convergence of the series for the conserved charges can be used to identify the delocalized and localized phases instead of the eigenfunctions.

5. Aubry-Andre model

Having constructed the conserved charges for a model with finite-range hopping and defined a condition for convergence of the power series for them, we can further investigate the meaning of our convergence criterion. In particular, since our goal is to identify the validity of our construction of the conservation laws with the presence of localization, the power series should fail to converge according to our criterion in a delocalized phase.

We thus require a non-interacting model with disorder in one dimension which has a delocalized phase. While any model with finite-range hopping and an on-site random potential in one dimension always produces localization [1, 3], a quasi-periodic potential can produce localized and delocalized phases. Such a model is the Aubry-Andre model [42] given by the Hamiltonian

$$H = h\sum_j \cos(2\beta\pi j)c_j^\dagger c_j - \frac{1}{2}\sum_j (c_j^\dagger c_{j+1} + \text{h.c.}), \tag{39}$$

where β is an irrational number. The parameter h can be tuned to effect a transition from a localized phase (for $h > 1$) to a delocalized phase (for $h < 1$) [42]. We note that this model is usually studied with an additional term that introduces a p-wave pairing gap [54], but we set it equal to zero for our analysis.

The localized phase here is one in which *all* single particle states are localized and similarly *all* single particle states are delocalized in the delocalized phase. The transition between these phases happens at $h = 1$. Since the Hamiltonian in Eq. (39) is also of the form (18), we can use the expressions obtained for the η_{ij}^m in the previous section to construct the conserved charges. These will now depend on the parameter h (i.e. $y \to (2h)^{-1}$ in the previous section) and if the criterion for convergence postulated by us is a valid one to detect localization, we should observe the power series to converge in the localized phase ($h > 1$) and diverge in the delocalized phase ($h < 1$). This is indeed the case as we see e.g. from Fig. 5, which shows that a typical matrix element of η^m goes to zero quite rapidly with increasing m for $h > 1$ but diverges for $h < 1$. Thus, we have established that our convergence criterion is valid for identifying the localization–delocalization transition.

6. Interactions

We now turn to systems with interactions. The simplest way to introduce interactions to models we studied here is through a nearest neighbor density-density term. Let us,

for example, add such a term to Eq. (18),

$$\begin{aligned} H &= \sum_i \epsilon_i n_i - ty \sum_i (c_i^\dagger c_{i+1} + h.c.) + V \sum_i n_i n_{i+1} \\ &= H_0 + V\delta H, \end{aligned} \tag{40}$$

where we redefined H_0 as compared to Eq. (18).

We assume that the particles here are spineless fermions. It is tempting to try a construction of the conserved charges starting from a zeroth order Hamiltonian that combines the on-site and interaction terms since they commute with each other and their eigenstates are localized at every site. However, the interaction term is quartic in creation and annihilation operators and so the conserved charges can no longer be assumed to be power series in the hopping with each term quadratic in the creation and annihilation operators. Such an assumption leads to no solution for the coefficients since the commutators keep producing terms with increasingly longer trails of creation and annihilation operators as one goes to higher orders in the hopping. A more profitable exercise is to try to obtain the conserved charges as power series in the hopping but only to the first order in the interaction. While these are not exact, they offer a reasonable approximation in the limit of small interaction strength. Weak interactions typically should not destroy the localization present in the non-interacting limit and thus conserved charges should continue to exist.

We know from our previous calculation that the operator of the form $Q_0 = n_0 + \sum_{ijm} \eta_{ij}^m y^m c_i^\dagger c_j$ commutes with H_0. Let us now define a new operator $Q = Q_0 + V\delta Q$ to linear order in V and calculate the commutator.

$$\begin{aligned} [Q, H] &= [Q_0 + V\delta Q, H_0 + V\delta H] \\ &= V\left([\delta Q, H_0] + [Q_0, \delta H]\right) + O(V^2). \end{aligned} \tag{41}$$

We choose δQ such that $[\delta Q, H_0] + [Q_0, \delta H] = 0$, so that Q and H commute to $O(V)$. We assume the form $\delta Q = \sum_{rstv} \psi_{rstv} c_r^\dagger c_s c_t^\dagger c_v$. Note this is quartic in the creation and annihilation operators since the interaction term is as well. Thus,

$$\begin{aligned} [\delta Q, H_0] + [Q_0, \delta H] = 0 \\ \sum_{kim} y^m (\eta_{ik}^m c_i^\dagger c_k n_{k+1} - \eta_{ki}^m c_k^\dagger c_i n_{k+1} + \eta_{i,k+1}^m n_k c_i^\dagger c_{k+1} \\ -\eta_{k+1,i}^m n_k c_{k+1}^\dagger c_i) = ty \sum_{rstv} (\psi_{rstv-1} - \psi_{rst-1v} + \psi_{rstv+1} - \\ \psi_{rst-1v} + \psi_{rs-1tv} - \psi_{r+1stv} + \psi_{rs+1tv} - \psi_{r-1stv}) c_r^\dagger c_s c_t^\dagger c_v \\ - \sum_{rstv} \psi_{rstv} (\epsilon_v - \epsilon_t + \epsilon_s - \epsilon_r) c_r^\dagger c_s c_t^\dagger c_v. \end{aligned}$$

We now assume that ψ_{rstv} can be written as a power series in y, i.e. $\psi_{rstv} = \sum_\alpha A_\alpha^{rstv} y^\alpha$. Equating the coefficients at different orders of y, one can in principle obtain A_m^{rstv} in terms of the η_{ij}^m for the case with $V = 0$. In fact, it can be seen that at a given order m, the A_m^{rstv} are linear combinations of the η_{ij}^m and the A_{m-1}^{rstv}. One can

also impose constraints arising from the anti-commutation of the fermionic operators, the Hermitian nature of the conservation laws and the number of non-zero components of the $\eta_{ij}^{(m)}$ to severely constrain the number of non-zero components of A_m^{rstv}.

Let us, for example, derive δQ to the first order in y, i.e. we set $m = 1$. We have

$$\sum_{ki}(\eta_{ik}^1 c_i^\dagger c_k n_{k+1} - \eta_{ki}^1 c_k^\dagger c_i n_{k+1} + \eta_{i,k+1}^1 n_k c_i^\dagger c_{k+1} - \eta_{k+1,i}^1 n_k c_{k+1}^\dagger c_i) = \sum_{rstv} A_1^{rstv}(-\epsilon_v + \epsilon_t - \epsilon_s + \epsilon_r) c_r^\dagger c_s c_t^\dagger c_v.$$

Since, only $\eta_{i,i+1}^1$ and $\eta_{i,i-1}^1$ are non-zero, the non-zero A_1^{rstv} are given by the following equations:

$$\begin{aligned}
A_1^{k+1,k,k+1,k+1} = A_1^{k,k+1,k+1,k+1} &= \frac{\eta_{k+1,k}^1}{\epsilon_{k+1} - \epsilon_k} \\
A_1^{k,k-1,k+1,k+1} = A_1^{k-1,k,k+1,k+1} &= \frac{\eta_{k-1,k}^1}{\epsilon_{k-1} - \epsilon_k} \\
A_1^{k,k,k+1,k} = A_1^{k,k,k,k+1} &= \frac{\eta_{k,k+1}^1}{\epsilon_k - \epsilon_{k+1}} \\
A_1^{k,k,k+2,k+1} = A_1^{k,k,k+1,k+2} &= \frac{\eta_{k+2,k+1}^1}{\epsilon_{k+2} - \epsilon_{k+1}}.
\end{aligned}$$

The corresponding expression for δQ to order y is

$$\begin{aligned}
\delta Q = yV \sum_{rstv} A_1^{rstv} c_r^\dagger c_s c_t^\dagger c_v = yV \sum_k \Bigg[&\frac{\eta_{k+1,k}^1}{\epsilon_{k+1} - \epsilon_k}(c_{k+1}^\dagger c_k \\
&+ c_k^\dagger c_{k+1}) n_{k+1} + \frac{\eta_{k-1,k}^1}{\epsilon_{k-1} - \epsilon_k}(c_{k-1}^\dagger c_k + c_k^\dagger c_{k-1}) n_{k+1} \\
&+ \frac{\eta_{k,k+1}^1}{\epsilon_k - \epsilon_{k+1}} n_k (c_{k+1}^\dagger c_k + c_k^\dagger c_{k+1}) + \frac{\eta_{k+2,k+1}^1}{\epsilon_{k+2} - \epsilon_{k+1}} n_k \\
&(c_{k+2}^\dagger c_{k+1} + c_{k+1}^\dagger c_{k+2}) \Bigg].
\end{aligned}$$

Other approaches to construct conservation laws for interacting systems have been proposed including a recent one where the interacting problem is mapped onto a non-Hermitian problem on a lattice in operator space [53]. A convergence criterion for the resultant series based on the operator norm is then used to identify localized and delocalized phases.

7. Conclusions and discussion

Inspired by the Type-1 Hamiltonian system, we have demonstrated a scheme to obtain the conserved charges for non-interacting disordered models displaying localization in one dimension. One of our motivation was an observation of similarities between

localized and integrable systems, such as the absence of level repulsion and the absence of thermalization. Our conserved charges are exhibited as a power series in the hopping, and using a suitable convergence criterion, we show that the convergence (or divergence) of conserved charges tracks the presence (or absence) of localization. An interesting issue of "gauge dependence" of the conserved charges is unearthed and explored. It is shown that a full understanding of the gauge dependence leads to considerable simplifications of the charges in some cases. On the other hand, straightforward Rayleigh Schrödinger perturbation theory or equivalent schemes, commit one to a particular gauge that is often inconvenient.

This work provides a novel link between the concepts of localization and integrability. Our results hold within the context of the 1-d Anderson model, where all states are localized, and the Andre-Aubry model, where (all) states undergo a transition tuned by a coupling constant. It is not immediately obvious how to extend these results to a higher dimensional Anderson model with a mobility edge separating the two classes of states. The Aubry-Andre model exhibits an interesting kind of duality which allows the localized and delocalized phases to be mapped onto each other with the roles of the hopping and onsite potential exchanged. The duality transformation is expressed in terms of new fermonic operators given by $c_{\bar{k}} = \frac{1}{\sqrt{L}} \sum_n \exp(i2\pi\bar{k}\beta)c_n$, which are eigenstates of the momentum operator with eigenvalue: $k = \bar{k}F_{n-1} \mod F_n$, where F_n is the n-th Fibonacci number and $L = F_n$ [55, 56]. In terms of these fermionic operators the Hamiltonian (39) becomes

$$\frac{H}{h} = \frac{1}{h}\sum_{\bar{k}} \cos(2\beta\pi\bar{k}) n_{\bar{k}} - \frac{1}{2}\sum_{\bar{k}} (c^\dagger_{\bar{k}} c_{\bar{k}+1} + \text{h.c.}). \tag{42}$$

The Hamiltonian satisfies the duality relation: $H(h)/h = H(1/h)$. We have shown that for the Aubrey-Andre Hamiltonian written in real space, one can construct set of conserved charges that converge for $0 < h < 1$. Because of the duality of the model one can construct similar conserved charges in terms of $c_{\bar{k}}$ and $c^\dagger_{\bar{k}}$. The power series of these charges converge when $0 < 1/h < 1$ and both sets of charges diverge at $h = 1$. Thus, the duality of the model allows us to explicitly construct conservation charges in one phase given that they exist in the other.

This can be better understood by noting that localization is a basis dependent concept. We have been using localization (as is the standard practice) to mean localization in real space. To obtain the conserved charges for such a localized phase, we start from a Hamiltonian whose eigenstates are perfectly localized in real space and then add terms perturbatively in the hopping. Similarly, the delocalized phase of the Aubry-Andre model is localized in momentum space and one can then obtain its conserved charges by starting with a Hamiltonian perfectly localized in momentum space (tight binding model) and then add terms perturbatively in the on-site potential. This is the essence of the duality outlined above. Thus, the conserved charges also carry labels indicating the space (real or momentum) where the system is localized. What is important though is that once the basis in which the system is localized is

identified and the conserved charges are constructed accordingly, they are sensitive to the onset of delocalization in that basis and can be used to locate localization-delocalization transitions.

The importance of the basis can be further understood when one compares the behavior hard-core bosons with that of spinless fermions in the Aubry-Andre model [57, 58]. The duality between the localized and delocalized phases is destroyed for hard-core bosons. As a result, the relaxation of real space local observables in the localized phase is different from their conjugates in momentum space in the delocalized phase. This feature is absent for spinless fermions where the duality holds and as a consequence, conserved charges of the type derived in this work exist in both phases.

While it is only possible to construct these charges to lowest order in the interaction using our procedure, their fate upon the introduction of interactions can in principle be investigated numerically, which we defer to a future work.[+]

8. Acknowledgements

The work at UCSC was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES) under Award # FG02-06ER46319.. E. A. Y. was financially supported in part by the David and Lucile Packard Foundation. RM acknowledges support from the UGC-BSR Fellowship. SM thanks the DST, Government of India and the India-Israel joint research programme for funding.

References

[1] Anderson P W 1958 *Phys. Rev.* **109** 1492
[2] Lee P A and Ramakrishnan T 1985 *Rev. Mod. Phys.* **57** 287
[3] Abrahams E, Anderson P, Licciardello D and Ramakrishnan T 1979 *Phys. Rev. Lett.* **42** 673–676
[4] Deutsch J M 1991 *Phys. Rev. A* **43** 2046–2049
[5] Srednicki M 1994 *Phys. Rev. E* **50** 889–901
[6] Rigol M, Dunjko V, Yurovsky V and Olshanii M 2007 *Phys. Rev. Lett.* **98** 050405
[7] Basko D, Aleiner I and Altshuler B 2006 *Ann. Phys.* **321** 1126–1205
[8] Huse D and Nandkishore R 2015 *Annu. Rev. Cond. Matt. Phys.* **6**
[9] Pal A and Huse D A 2010 *Phys. Rev. B* **82** 174411
[10] Vosk R and Altman E 2013 *Phys. Rev. Lett.* **110** 067204
[11] Oganesyan V and Huse D A 2007 *Phys. Rev. B* **75** 155111
[12] Bardarson J H, Pollmann F and Moore J E 2012 *Phys. Rev. Lett.* **109** 017202
[13] Agarwal K, Gopalakrishnan S, Knap M, Müller M and Demler E 2014 *arXiv preprint arXiv:1408.3413*
[14] Shastry B S 1986 *Phys. Rev. Lett.* **56** 2453–2455
[15] Shastry B S 1986 *Phys. Rev. Lett.* **56**(15) 1529–1531
[16] Lüscher M 1976 *Nucl. Phys. B* **117**
[17] Grosse H 1989 *Lett. Math. Phys.* **18** 151
[18] Grabowski M and Mathieu P 1995 *Ann. Phys.* **243** 299–371
[19] Links J, Zhou H Q, McKenzie R H and Gould M D 2001 *Phys. Rev. Lett.* **86** 5096

[+] Note that the introduction of interactions destroys the duality of the model since it no longer has the same form in real and momentum space under the duality transform.

[20] Zhou H, Jiang L and Tang J 1990 *J. Phys. A* **23** 213
[21] Fuchssteiner B 1988 Symmetries and nonlinear phenomena (Singapore: World Scientific Publishers) pp 22–50
[22] Relano A, Dukelsky J, Gomez J M G and Retamosa J 2004 *Phys. Rev. E* **70** 026208
[23] Sklyanin E 1989 *J. Sov. Math.* **47** 2473
[24] Caux J S and Mossel J 2011 *J. Stat. Mech.* P02023
[25] Yuzbashyan E A and Shastry B S 2013 *J. Stat. Phys.* **150** 704–721
[26] Shastry B S 2005 *J. Phys. A: Math. Theor.* **38** L431
[27] Owusu H K, Wagh K and Yuzbashyan E A 2009 *J. Phys. A: Math. Theor.* **42** 035206
[28] Shastry B S 2011 *J. Phys. A: Math. Theor.* **44** 052001
[29] Owusu H K and Yuzbashyan E A 2011 *J. Phys. A* **44** 395302
[30] Owusu H and Shastry B S 2013 *J. Phys. A: Math. Theor.* **46** 375003
[31] Rigol M 2009 *Phys. Rev. Lett.* **103** 100403
[32] Berry M V and Tabor M 1977 *Proc. R. Soc. London, Ser. A* **356** 375–394
[33] Modak R, Mukerjee S and Ramaswamy S 2014 *Phys. Rev. B* **90** 075152
[34] Modak R and Mukerjee S 2014 *New J. Phys.* **16** 093016
[35] Rabson D A, Narozhny B and Millis A 2004 *Phys. Rev. B* **69** 054403
[36] Huse D A, Nandkishore R and Oganesyan V 2014 *Phys. Rev. B* **90** 174202
[37] Serbyn M, Papić Z and Abanin D A 2013 *Phys. Rev. lett.* **111** 127201
[38] Chandran A, Kim I H, Vidal G and Abanin D A 2015 *Phys. Rev. B* **91** 085425
[39] Vosk R, Huse D A and Altman E 2014 *arXiv preprint arXiv:1412.3117*
[40] Weaire D and Williams A 1977 *J. Phys. C* **10** 1239
[41] Johri S and Bhatt R 2014 *Phys. Rev. B* **90** 060205
[42] Aubry S and André G 1980 *Ann. Israel Phys. Soc* **3** 18
[43] Ossipov A 2013 *J. Phys. A: Math. Theor.* **46** 105001
[44] Nishino S, Matsuda H and Goda M 2007 *J. Phys. Soc. Jpn.* **76**
[45] Chalker J, Pickles T and Shukla P 2010 *Phys. Rev. B* **82** 104209
[46] Yuzbashyan E A, Baytin A A and Altshuler B L 2005 *Phys. Rev. B* **71** 094505
[47] von Delft J 2001 *Annalen der Physik* **3** 219
[48] Anderson P W 1959 *J. Phys. Chem. Solids* **11** 26
[49] Schechter M, von Delft J, Imry Y and Levinson Y 2003 *Phys. Rev. B* **67** 064506
[50] Mehta M L 2004 *Random matrices* vol 142 (Academic press)
[51] Imbrie J Z 2014 *arXiv preprint arXiv:1403.7837*
[52] Mukerjee S, Oganesyan V and Huse D 2006 *Phys. Rev. B* **73** 035113
[53] Ros V, Müller M and Scardicchio A 2015 *Nucl. Phys. B* **891** 420–465
[54] Cai X, Lang L J, Chen S and Wang Y 2013 *Phys. Rev. Lett.* **110** 176403
[55] Roósz G m H, Divakaran U, Rieger H and Iglói F 2014 *Phys. Rev. B* **90**(18) 184202
[56] Iyer S, Oganesyan V, Refael G and Huse D A 2013 *Phys. Rev. B* **87** 134202
[57] Gramsch C and Rigol M 2012 *Phys. Rev. A* **86** 053615
[58] He K, Santos L F, Wright T M and Rigol M 2013 *Phys. Rev. A* **87** 063637